

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 16, 2003, 12:12:16 ; Search time 12 seconds
(without alignments)
411.307 Million cell updates/sec

Title: US-10-015-967-2
Perfect score: 644
Sequence: 1 MKVLISLLPLPLMSMV.....SRACQFLKQCQLRSFALPL 119

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	77	12.0	2269 1	RRPL_SV41
2	74.5	11.6	448 1	VDR_COTUA
3	74.5	11.6	451 1	VDR_CHICK
4	72.5	11.3	3313 1	CLR3_RAT
5	71	11.0	411 1	SV44_BRARE
6	71	11.0	1798 1	LMB2_HUMAN
7	71	11.0	2262 1	RRPL_P12HT
8	70	10.9	343 1	70P1_DRONE
9	70	10.9	746 1	70P1_DRONE
10	69.5	10.8	575 1	TRBM_HUMAN
11	69	10.7	314 1	SIX3_CHICK
12	68.5	10.6	422 1	VDR_MOUSE
13	68.5	10.6	422 1	VDR_XENLA
14	68.5	10.6	423 1	VDR_XENLA
15	68.5	10.6	424 1	VDR_BOVIN
16	68.5	10.6	427 1	VDR_HUMAN
17	68.5	10.6	647 1	VE1_HPV40
18	67	10.4	475 1	HUTM_BACSD
19	67	10.4	529 1	VE98_HUMAN
20	66.5	10.3	422 1	COT1_MOUSE
21	66.5	10.3	423 1	COT1_HUMAN
22	66.5	10.3	424 1	COT1_BOVIN
23	66.5	10.3	3301 1	CLR3_MOUSE
24	66	10.2	1539 1	SMCY_HUMAN
25	65.5	10.2	646 1	VE1_HPV07
26	65.5	10.2	706 1	FZD6_HUMAN
27	65	10.1	363 1	PELI_COLLN
28	65	10.1	441 1	COAT_SOCMV
29	64.5	10.0	622 1	MAK_MOUSE
30	63.5	9.9	262 1	YOS2_CAEEL
31	63.5	9.9	326 1	Z186_HUMAN
32	63	9.8	160 1	BCP_PANTR
33	63	9.8	246 1	SIX6_HUMAN

34	63	9.8	332 1	SIX3_HUMAN	095343 homo sapien
35	63	9.8	333 1	SIX3_MOUSE	062233 mus musculu
36	63	9.8	371 1	RA14_YEAST	P28519 saccharomyc
37	62.5	9.7	156 1	RNR7_HUMAN	091el homo sapien
38	62.5	9.7	410 1	COT2_CHICK	090733 gallus gall
39	62.5	9.7	414 1	COT2_BOVIN	P24468 homo sapien
40	62.5	9.7	414 1	COT2_HUMAN	P43135 mus musculu
41	62.5	9.7	414 1	COT2_MOUSE	009018 rattus norv
42	62.5	9.7	414 1	COT2_RAT	P12687 saccharomyc
43	62	9.6	371 1	RNR2_YEAST	004913 xenopus lae
44	62	9.6	577 1	NR42_XENLA	001955 homo sapien
45	62	9.6	1670 1	CA34_HUMAN	

ALIGNMENTS

```

RESULT 1
ID RRPL_SV41          STANDARD;          PRT; 2269 AA.
AC P35341;
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE RNA polymerase beta subunit (EC 2.7.7.48) (Large structural protein)
DE (L protein).
GN L.
OS Simian virus 41 (SV41).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Rubulaviruses.
OX NCBI_TaxID=11228;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Toshiya/Chanock;
RX MEDLINE=93019033; PUBMED=1328485;
RA Ogawa M., Mutsuga N., Tsuridome M., Kawano M., Matsumura H.,
RA Kusagawa S., Komada H., Nishio M., Ito Y.;
RT "Nucleotide sequence analysis of the simian virus 41 gene encoding
RT the large (L) protein and construction of a phylogenetic tree for the
RT L proteins of paramyxoviruses."
RL J. Gen. Virol. 73:2743-2750(1992).
CC - FUNCTION: PROBABLE COMPONENT OF THE ACTIVE POLYMERASE. IT MAY
CC FUNCTION IN MRNA SYNTHESIS, CAPING, METHYLATION AND POLY(A)
CC SYNTHESIS OF NEWLY SYNTHESIZED VIRAL MRNAs, RNA EDITING OF THE P
CC GENE TRANSCRIPT. AND PROTEIN KINASE ACTIVITY.
CC - CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC (RNA)(N).
CC - SIMILARITY: WITH L PROTEIN OF OTHER PARAMYXOVIRUSES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
CC -----
CC EMBL: X64375; CAA45569.1; -.
CC PIR: J01750; J01750.
CC InterPro: IPR001016; Viral_RNA_pol_L.
CC Pfam: PF00946; Paramyx_RNA_pol; 1.
CC Trnseqase: RNA-directed RNA polymerase.
CC SEQUENCE 2269 AA; 256429 MW; 3BD60C14AA1615B CRC64;
CC -----
Query Match 12.0%; Score 77; DB 1; Length 2269;
Best Local Similarity 27.5%; Pred. No. 3.8;
Matches 30; Conservative 12; Mismatches 31; Indels 36; Gaps 6;
OY 13 PLMLMSVSSSLNPVAGHNRGQAS-----RRRLQ-----EGQDEE--CK 53
DB 422 PLPLGNAKSKL---IELHDNDSISYEVYLRHKKSLLEFKKCFDFPDESLTFMK 477
OY 54 DWFLAPRRKFTVSGLPKQCPDFHFGNVKKTGRHRRH--RRPKNSR 101

```


FT 2N_PING 47 67 C4-TYPE.
 FT 2N_PING 83 107 C4-TYPE.
 FT DOMAIN 113 215 HINGE.
 FT DOMAIN 216 451 LIGAND-BINDING.
 SQ SEQUENCE 451 AA; 51299 MW; 2078B6A6C8DBE5FC CRC64;
 Query Match 11.6%; Score 74.5; DB 1; Length 451;
 Best Local Similarity 32.9%; Pred. No. 1.5;
 Matches 24; Conservative 11; Mismatches 15; Indels 23; Gaps 6;
 QY 50 CE-CRMEFLAPRRKRWTSGLPKKQPCDHFENKTRHGHKHPKHSACQFLK 108
 DB 64 CECCGKGFRRSMRKAMFT-----CP---FNGDKRTKDN-----RHCQACR--LK 105
 QY 109 QC-----QLRSEFL 117
 DB 106 RCYDGMKEFIL 118
 RESULT 4
 CLR3 RAT STANDARD; PRT; 3313 AA.
 AC 088278;
 DT 15-JUN-2002 (rel. 41, Created)
 DT 15-JUN-2002 (rel. 41, Last sequence update)
 DT 15-JUN-2002 (rel. 41, Last annotation update)
 DE Cadherin EGF LAG seven-pass G-type receptor 3 precursor (Multiple
 DE epidermal growth factor-like domains 2).
 GN CELSR3 OR MEGR2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Rattus.
 RX NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=98360089; PubMed=9693030;
 RA Nakayama M., Nakajima D., Nagase T., Nomura N., Seki N., Ohara O.;
 RT Identification of high-molecular-weight proteins with multiple
 RT EGF-like motifs by motif-trap screening.";
 RL Genomics 51:27-34(1998).
 CC -1- FUNCTION: Receptor that may have an important role in cell/cell
 CC signaling during nervous system formation.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Expressed in the brain. Expressed in
 CC cerebellum, olfactory bulb, cerebral cortex, hippocampus and
 CC brain stem.
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED RECEPTORS.
 CC -1- SIMILARITY: CONTAINS 9 CADHERIN DOMAINS.
 CC -1- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ G-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ EGF-LIKE DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 GPS DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB011528; BAA32459.1; -
 DR HSSP; P00740; 1EDM.
 DR InterPro: IPR000152; ASX_hydroxyl.
 DR InterPro: IPR002126; Cadherin.
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000742; EGF_2.
 DR InterPro: IPR001881; EGF_Ca.
 DR InterPro: IPR000832; GPCR_secretin.
 DR InterPro: IPR001879; hormn_receptor.
 DR InterPro: IPR002049; Laminin_EGF.
 DR InterPro: IPR001791; Laminin_G.

DR InterPro: IPR00203; PKD_cys_rich.
 DR Pfam: PF00002; 7tm_2; 1.
 DR Pfam: PF00028; cadherin; 9.
 DR Pfam: PF00008; EGF; 6.
 DR Pfam: PF01825; GPS; 1.
 DR Pfam: PF02793; HRM; 1.
 DR Pfam: PF00054; Laminin_G; 1.
 DR PRINTS: PR00205; CADHERIN.
 DR SMART: SM00112; CA; 8.
 DR SMART: SM00180; EGF_Lam; 1.
 DR SMART: SM00001; EGF_like; 6.
 DR SMART: SM00303; GPS; 1.
 DR SMART: SM00008; Hormr; 1.
 DR SMART: SM00282; Lamg; 2.
 DR PROSITE: PS00010; ASX_HYDROXYL; 1.
 DR PROSITE: PS00232; CADHERIN_1; 7.
 DR PROSITE: PS50268; CADHERIN_2; 8.
 DR PROSITE: PS00022; EGF_1; 6.
 DR PROSITE: PS01186; EGF_2; 4.
 DR PROSITE: PS50221; GPS; 1.
 DR PROSITE: PS50025; LAM_G_DOMAIN; 2.
 DR PROSITE: PS00649; G_PROTEIN_RECP_F2_1; FALSE_NEG.
 DR PROSITE: PS50227; G_PROTEIN_RECP_F2_3; 1.
 DR PROSITE: PS50261; G_PROTEIN_RECP_F2_4; 1.
 DR PROSITE: PS01248; LAMININ_TYPE_EGF; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW EGF-like domain; Calcium-binding; Laminin EGF-like domain; Repeat;
 KW Developmental protein; Hydroxylation; Signal.
 FT CHAIN 1 31
 FT SIGNAL 32 3313
 FT TRANSMEM 2538 2538
 FT TRANSMEM 2539 2559
 FT DOMAIN 2560 2570
 FT TRANSMEM 2571 2591
 FT DOMAIN 2592 2599
 FT TRANSMEM 2600 2620
 FT DOMAIN 2621 2641
 FT TRANSMEM 2642 2662
 FT TRANSMEM 2663 2679
 FT TRANSMEM 2680 2700
 FT TRANSMEM 2701 2724
 FT TRANSMEM 2725 2745
 FT TRANSMEM 2746 2752
 FT TRANSMEM 2753 2773
 FT DOMAIN 2774 3313
 FT DOMAIN 317 424
 FT DOMAIN 425 536
 FT DOMAIN 537 642
 FT DOMAIN 643 747
 FT DOMAIN 748 849
 FT DOMAIN 850 952
 FT DOMAIN 953 1058
 FT DOMAIN 1059 1160
 FT DOMAIN 1059 1257
 FT DOMAIN 1366 1424
 FT DOMAIN 1426 1462
 FT DOMAIN 1466 1505
 FT DOMAIN 1506 1505
 FT DOMAIN 1506 1710
 FT DOMAIN 1713 1749
 FT DOMAIN 1753 1935
 FT DOMAIN 1937 1972
 FT DOMAIN 1973 2011
 FT DOMAIN 2012 2044
 FT DOMAIN 2046 2081
 FT DOMAIN 2087 2120
 FT DOMAIN 2475 2527
 FT DISULFID 1370 1381
 FT DISULFID 1375 1412
 FT DISULFID 1414 1423
 FT DISULFID 1430 1441
 FT DISULFID 1435 1450
 1 (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 2 (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 3 (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 4 (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 5 (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 6 (POTENTIAL).
 EXTRACELLULAR (POTENTIAL).
 7 (POTENTIAL).
 CYTOPLASMIC (POTENTIAL).
 CADHERIN 1.
 CADHERIN 2.
 CADHERIN 3.
 CADHERIN 4.
 CADHERIN 5.
 CADHERIN 6.
 CADHERIN 7.
 CADHERIN 8.
 CADHERIN 9.
 EGF-LIKE 1.
 EGF-LIKE 2.
 EGF-LIKE 3.
 EGF-LIKE 4.
 EGF-LIKE 5.
 EGF-LIKE 6.
 EGF-LIKE 7.
 EGF-LIKE 8.
 EGF-LIKE 9.
 LAMININ G-LIKE 1.
 LAMININ G-LIKE 2.
 LAMININ G-LIKE 3.
 LAMININ G-LIKE 4.
 LAMININ G-LIKE 5.
 LAMININ G-LIKE 6.
 LAMININ G-LIKE 7.
 LAMININ G-LIKE 8.
 LAMININ G-LIKE 9.
 GPS.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.
 BY SIMILARITY.

```

FT DISULFID 1452 1461 BY SIMILARITY.
FT DISULFID 1470 1481 BY SIMILARITY.
FT DISULFID 1475 1491 BY SIMILARITY.
FT DISULFID 1493 1504 BY SIMILARITY.
FT DISULFID 1717 1728 BY SIMILARITY.
FT DISULFID 1722 1737 BY SIMILARITY.
FT DISULFID 1739 1746 BY SIMILARITY.
FT DISULFID 1941 1952 BY SIMILARITY.
FT DISULFID 1946 1961 BY SIMILARITY.
FT DISULFID 1963 1972 BY SIMILARITY.
FT DISULFID 1976 1987 BY SIMILARITY.
FT DISULFID 1981 1999 BY SIMILARITY.
FT DISULFID 2001 2010 BY SIMILARITY.
FT DISULFID 2010 2028 BY SIMILARITY.
FT DISULFID 2018 2031 BY SIMILARITY.
FT DISULFID 2033 2043 BY SIMILARITY.
FT DISULFID 2050 2065 BY SIMILARITY.
FT DISULFID 2052 2068 BY SIMILARITY.
FT DISULFID 2070 2080 BY SIMILARITY.
FT MOD_RES 1954 1954 HYDROXYLATION (POTENTIAL).
FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 838 838 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1173 1173 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1213 1213 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1308 1308 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1318 1318 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1640 1640 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1704 1704 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1761 1761 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2044 2044 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2173 2173 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2192 2192 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2382 2382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2472 2472 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 2504 2504 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 3313 AA; 359348 MW; B11DA09517288764 CRC64;

```

```

Query Match 11.3%; Score 72.5; DB 1; Length 3313;
Best Local Similarity 27.6%; Pred. No. 16;
Matches 35; Conservative 9; Mismatches 40; Indels 43; Gaps 7;

```

```

OY 24 LNDVGARHGRDQASRRMLQEG--GOECCCKMFL-----RAPRRKFTVSGLP 71
DB 1978 LND-----COMGSGCRHLQSGPHGYTCDCASGQHCENRMDQCCPRGMMGSPICGP 2030
OY 72 KKOCPDHFHKG---NYKTRHQRH---HRRP-----NKHSRACQQLKOC 110
DB 2031 ---CNCDDVHKFPDPCNKTKSGGCHCKEHPHYPAGSDCLPCDCYVSGTSTRSCAPHSQC 2087
OY 111 QLRSPAL 117
DB 2088 PCRPGAL 2094

```

```

RESULT 5
SV44_BRARE STANDARD; PRT; 411 AA.
ID SV44_BRARE
AC Q06725;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Steroid receptor homolog SVP 44.
GN SVP44 OR NR2F1.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93223680; PubMed=8467797;
RA Fjose A., Norres S., Weber U., Mlodzik M.;
RT "Functional conservation of vertebrate seven-up related genes in

```

```

RT neurogenesis and eye development.";
RL EMBL J. 12:1403-1414(1993).
CC -1- FUNCTION: PUTATIVE RECEPTOR THAT IS REQUIRED IN PHOTORECEPTOR
CC CELLS PRECURSORS DURING EYE DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC NR2 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X70299; CAA49780.1; -
DR HSSP: P19793; 2NUL.
DR ZFIN: ZDB-GENE-980526-115; nr2f1.
DR InterPro: IPR000536; Hormone_rec_11g.
DR InterPro: IPR001723; Stdrhm_receptor.
DR InterPro: IPR001628; Znfc4steroid.
DR Pfam: PF00104; hormone_rec; 1.
DR Pfam: PF00105; zf-c4; 1.
DR PRINTS: PR00398; STRDHOMONER.
DR PRINTS: PR00047; STROIDFTNGER.
DR ProDom: PD000035; Znfc4steroid; 1.
DR SMART: SM00430; HOL1; 1.
DR SMART: SM00399; Znfc4; 1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
DR Receptor: Transcription regulation; DNA-binding; Nuclear protein;
DR Zinc-finger; Vision.
KW DNA_BIND 76 141 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 76 96 C4-TYPE.
FT ZN_FING 112 136 C4-TYPE.
SQ SEQUENCE 411 AA; 45482 MW; B0405F4AEFA0202 CRC64;

```

```

Query Match 11.0%; Score 71; DB 1; Length 411;
Best Local Similarity 26.3%; Pred. No. 3.2;
Matches 26; Conservative 14; Mismatches 27; Indels 32; Gaps 6;

```

```

OY 26 PGVA--RGHRDQASRR-----WLDGGGCE--CKDWIIRADRRKFTVSGLP 71
DB 56 PGTAGDKSGNSGSOQHIEVCVCGDKSSGKHNGQFTCEGKSFKRSVRRN-LTYTCRA 114
OY 72 KKOCPDHFHKGNYKTRHQRHRRKPNKHSRACQQLKOC 110
DB 115 NRNCPID-----QHHRNCCOYCR-----LKKC 136

```

```

RESULT 6
LMB2_HUMAN STANDARD; PRT; 1798 AA.
ID LMB2_HUMAN
AC P55268; Q16321.
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Laminin beta-2 chain precursor (S-laminin) (Laminin B1s chain).
GN LAMB2 OR LAMS.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95213013; PubMed=7698745;
RA Wewer U.M., Gerecke D.R., Durkin M.E., Kurtz K.S., Mattei M.-G.,
RA Champilaud M.F., Burgeson R.E., Albrechtsen R.;
RT "Human beta 2 chain of laminin (formerly S chain): cDNA cloning,
RT chromosomal localization, and expression in carcinomas.";
RL Genomics 24:243-252(1994).
RN [2]
RP SEQUENCE FROM N.A.

```

RA MEDLINE-95316263; PubMed-7795887;
RA Iivanainen A., Voutilainen R., Sainio K., Eddy R., Shows T.B.,
RA Sariola H., Tryggvason K.,
RT "The human laminin beta 2 chain (S-laminin): structure, expression in
RT fetal tissues and chromosomal assignment of the LAMB2 gene.";
RL Matrix Biol. 14:489-497(1995).
CC -1- FUNCTION: Binding to cells via a high affinity receptor, laminin
CC is thought to mediate the attachment, migration, and organization
CC of cells into tissues during embryonic development by interacting
CC with other extracellular matrix components.
CC -1- SUBUNIT: Laminin is a complex glycoprotein, consisting of three
CC different polypeptide chains (alpha, beta, gamma), which are bound
CC to each other by disulfide bonds into a cross-shaped molecule
CC comprising one long and three short arms with globules at each
CC end.
CC THE BETA-2 CHAIN IS A SUBUNIT OF LAMININ-3 (S-LAMININ), LAMININ-4
CC (S-NEROSIN), AND LAMININ-7 (KS-LAMININ).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
CC COMPONENT). S-LAMININ IS CONCENTRATED IN THE SYNAPTIC
CC CLEFT OF THE NEUROSCULPTURAL JUNCTION.
CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC -1- DOMAIN: DOMAINS VI AND IV ARE GLOBULAR.
CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -1- SIMILARITY: CONTAINS 12.5 LAMININ EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; 268155; CAA92279.1; -
DR EMBL; 268156; CAA92279.1; JOINED.
DR EMBL; X79683; CAA56130.1; -
DR EMBL; S77512; AAB34682.2; -
DR HSSP; P02468; IKLO.
DR Gene; HGNC:6487; LAMB2.
DR MIM: 150325; -
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001886; LamNT.
DR InterPro: IPR002049; Laminin_EGF.
DR Pfam; PF000053; laminin_EGF; 13.
DR Pfam; PF000055; laminin_Nterm; 1.
DR PRINTS; PR00011; EGF_LAMININ.
DR PRODOM; PD002082; LamNT; 1.
DR SMART; SM00180; EGF_Lam; 11.
DR SMART; SM00136; LamNT; 1.
DR PROSITE; PS00022; EGF_1; 10.
DR PROSITE; PS01186; EGF_2; 2.
DR PROSITE; PS01248; LAMININ_TYPE_EGF; 12.
KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.
KW SIGNAL
FT CHAIN 1 32
FT DOMAIN 33 1798 LAMININ BETA-2 CHAIN.
FT DOMAIN 283 346 LAMININ N-TERMINAL (DOMAIN VI).
FT DOMAIN 347 409 LAMININ EGF-LIKE 1.
FT DOMAIN 410 469 LAMININ EGF-LIKE 2.
FT DOMAIN 470 521 LAMININ EGF-LIKE 3.
FT DOMAIN 522 552 LAMININ EGF-LIKE 4.
FT DOMAIN 553 781 LAMININ EGF-LIKE 5 (INCOMPLETE).
FT DOMAIN 783 830 LAMININ DOMAIN IV.
FT DOMAIN 831 876 LAMININ EGF-LIKE 6.
FT DOMAIN 877 926 LAMININ EGF-LIKE 7.
FT DOMAIN 927 985 LAMININ EGF-LIKE 8.
FT DOMAIN 986 1037 LAMININ EGF-LIKE 9.
FT DOMAIN 1038 1094 LAMININ EGF-LIKE 10.
FT DOMAIN 1095 1142 LAMININ EGF-LIKE 11.
FT DOMAIN 1143 1189 LAMININ EGF-LIKE 12.

FT DOMAIN 1143 1189 LAMININ EGF-LIKE 13.
FT DOMAIN 1190 1409 DOMAIN II.
FT DOMAIN 1410 1442 DOMAIN ALPHA.
FT DOMAIN 1443 1798 DOMAIN I.
FT DOMAIN 1472 1526 COILED COIL (POTENTIAL).
FT DOMAIN 1577 1790 COILED COIL (POTENTIAL).
FT DOMAIN 283 292 BY SIMILARITY.
FT DISULFID 285 310 BY SIMILARITY.
FT DISULFID 312 321 BY SIMILARITY.
FT DISULFID 324 344 BY SIMILARITY.
FT DISULFID 347 356 BY SIMILARITY.
FT DISULFID 349 374 BY SIMILARITY.
FT DISULFID 377 386 BY SIMILARITY.
FT DISULFID 389 407 BY SIMILARITY.
FT DISULFID 410 423 BY SIMILARITY.
FT DISULFID 412 438 BY SIMILARITY.
FT DISULFID 440 449 BY SIMILARITY.
FT DISULFID 452 467 BY SIMILARITY.
FT DISULFID 470 484 BY SIMILARITY.
FT DISULFID 472 491 BY SIMILARITY.
FT DISULFID 493 502 BY SIMILARITY.
FT DISULFID 505 519 BY SIMILARITY.
FT DISULFID 783 795 BY SIMILARITY.
FT DISULFID 785 802 BY SIMILARITY.
FT DISULFID 804 813 BY SIMILARITY.
FT DISULFID 816 828 BY SIMILARITY.
FT DISULFID 831 843 BY SIMILARITY.
FT DISULFID 833 850 BY SIMILARITY.
FT DISULFID 852 861 BY SIMILARITY.
FT DISULFID 864 874 BY SIMILARITY.
FT DISULFID 877 886 BY SIMILARITY.
FT DISULFID 893 893 BY SIMILARITY.
FT DISULFID 896 905 BY SIMILARITY.
FT DISULFID 908 924 BY SIMILARITY.
FT DISULFID 927 943 BY SIMILARITY.
FT DISULFID 929 954 BY SIMILARITY.
FT DISULFID 956 965 BY SIMILARITY.
FT DISULFID 968 983 BY SIMILARITY.
FT DISULFID 986 1000 BY SIMILARITY.
FT DISULFID 988 1007 BY SIMILARITY.
FT DISULFID 1010 1019 BY SIMILARITY.
FT DISULFID 1022 1035 BY SIMILARITY.
FT DISULFID 1095 1107 BY SIMILARITY.
FT DISULFID 1097 1114 BY SIMILARITY.
FT DISULFID 1116 1125 BY SIMILARITY.
FT DISULFID 1128 1140 BY SIMILARITY.
FT DISULFID 1143 1155 BY SIMILARITY.
FT DISULFID 1145 1162 BY SIMILARITY.
FT DISULFID 1164 1173 BY SIMILARITY.
FT DISULFID 1176 1187 BY SIMILARITY.
FT DISULFID 1190 1190 INTERCHAIN (PROBABLE).
FT DISULFID 1193 1193 INTERCHAIN (PROBABLE).
FT DISULFID 1797 1797 INTERCHAIN (PROBABLE).
FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1085 1085 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1249 1249 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1308 1308 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1348 1348 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1499 1499 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 914 914 R -> G (IN REF. 2).
FT CONFLICT 1179 1179 G -> A (IN REF. 2).
SQ SEQUENCE 1798 AA; 196079 MW; 9555CF5B24850CB7 CRC64;

Query Match 11.0%; Score 71; DB 1; Length 1798;
Best Local Similarity 28.4%; Pred. No. 13;
Matches 27; Conservative 6; Mismatches 46; Indels 16; Gaps 4;
OY 23 SLNPGVARGHRDR-----GQASRRWLOEGGOEE-CDWFLPAPRRKFTVSGLPKROCP 76
DB 873 SCRPVCGNHADBCNHTGACLCGRDHTGSEHCEKCIAGHRPRPLPY-----GGQCRPCP 928

```

OY 77 CDHFGKGVKTRHQRHHRKPNKHSRACQFLKCCQ 111
DB 929 CPEGPGS-----QRHFATSCQHDYSSQIVCHCR 957

RESULT 7
RRPL_P12HT STANDARD; PRT; 2262 AA.
ID RRPL_P12HT
AC P26676;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE RNA polymerase beta subunit (EC 2.7.7.48) (large structural protein)
(L protein).
GN L.
OS Human parainfluenza 2 virus (strain Toshiba) (PIV-2).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Rubulaviruses.
OX NCBI_TaxID=11214;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91252221; PubMed=1645865;
RA Kawano M., Okamoto K., Bando H., Kondo K., Tsurudome M., Komada H.,
RA Nishio M., Ito Y.;
RT "Characterizations of the human parainfluenza type 2 virus gene
RT encoding the L protein and the intergenic sequences."
RL Nucleic Acids Res. 19:2739-2746(1991).
CC -1- FUNCTION: PROBABLE COMPONENT OF THE ACTIVE POLYMERASE. IT MAY
CC SYNTHESIZE IN MRNA SYNTHESIS, CAPING, METHYLATION AND POLY(A)
CC GENE TRANSCRIPT, AND PROTEIN KINASE ACTIVITY.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate -> N diphosphate +
CC (RNA)(N).
CC -1- SIMILARITY: WITH L PROTEIN OF OTHER PARAMYXOVIRUSES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X57559; CAA0788.1; -
DR PIR; S16664; S16664.
DR InterPro: IPR001016; VIRAL_RNA_POL_L.
DR Pfam: PF00946; Paramyx_RNA_POL_1.
KW Transferase; RNA-directed RNA polymerase.
SQ SEQUENCE 2262 AA; 256380 MW; 82468303F4453B48 CRC64;

Query Match 11.0%; Score 71; DB 1; Length 2262;
Best Local Similarity 26.6%; Pred. No. 16;
Matches 29; Conservative 13; Mismatches 31; Indels 36; Gaps 6;

OY 13 PLMLMSVSSSLNPGVARGHRGQAS-----RRWLQ-----EGGQECG-CK 53
DB 421 PLILPFAKSKSL-----IEFQNDMAELSYETLKHMKKEISLIERKCFDDPGSELSIFMK 476
OY 54 DMFLAPRRKRPMTYVSLPKKQCPDFKGNVKKTRQHRH-RKPNHNR 101
DB 477 DKAISAPRSDMSV-----FRSLIKQRHQRHHPNPFNR 513

RESULT 8
7UP1_DROME STANDARD; PRT; 543 AA.
ID 7UP1_DROME
AC P16375; Q9VGB0;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Steroid receptor seven-up type 1.
GN SVP OR NR2F3 OR CG11502.
OS Drosophila melanogaster (Fruit fly).

CC CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
CC CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
CC CC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
CC CC NCBI_TaxID=7227;
CC CC [1]
CC CC SEQUENCE FROM N.A.
CC CC RX MEDLINE=90124631; PubMed=2105166;
CC CC RA Mlodzik M., Hiromi Y., Weber U., Goodman C.S., Rubin G.M.;
CC CC RT "The Drosophila seven-up gene, a member of the steroid receptor gene
CC CC superfamily, controls photoreceptor cell fates."
CC CC RL Cell 60:211-224(1990).
CC CC [2]
CC CC RP SEQUENCE FROM N.A.
CC CC RC SRAIN-Berkeley;
CC CC RX MEDLINE=20196006; PubMed=10731132;
CC CC RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
CC CC RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
CC CC RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
CC CC RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,
CC CC RA Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
CC CC RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
CC CC RA Adell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
CC CC RA Ballew R.A., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
CC CC RA Beekoy D., Botchan M.R., Bouck J., Brokstein P., Brottler P.,
CC CC RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
CC CC RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
CC CC RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
CC CC RA Dodson R., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
CC CC RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrelis S., Fleischmann W.,
CC CC RA Flosser A., Gong F., Gorell J.H., Gu Z., Guan P., Harris K.,
CC CC RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
CC CC RA Hostin D., Houston K.A., Howland T.U., Wei M.-H., Ileguam C.,
CC CC RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
CC CC RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
CC CC RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
CC CC RA Liu X., Matel B., McIntosh T.C., McLeod M.P., McPherson D.,
CC CC RA Merulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
CC CC RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
CC CC RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,
CC CC RA Palazzolo M., Pittman K.S., Pan S., Pollard J., Puri V., Reese M.G.,
CC CC RA Relvert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
CC CC RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
CC CC RA Spier E., Spredling A.C., Stapleton M., Strong R., Sun E.,
CC CC RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
CC CC RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
CC CC RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
CC CC RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
CC CC RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
CC CC RT "The genome sequence of Drosophila melanogaster."
CC CC RL Science 287:2185-2195(2000).
CC CC -1- FUNCTION: PUTATIVE RECEPTOR THAT IS REQUIRED IN PHOTORECEPTOR
CC CC CELLS PRECURSORS DURING EYE DEVELOPMENT.
CC CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: TYPE 1 (SHOWN HERE) AND TYPE 2
CC CC (AC P16376); ARE PRODUCED BY ALTERNATIVE SPLICING AND ONLY DIFFER
CC CC IN THEIR C-TERMINAL PART.
CC CC -1- TISSUE SPECIFICITY: EXPRESSED IN A SUBSET OF NEURAL PRECURSORS.
CC CC -1- SIMILARITY: BEYONDS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC CC NR3 SUBFAMILY.
CC CC
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC
DR EMBL; M28863; AAA62770.1; -
DR EMBL; AE003695; AAF54773.1; -

```

DR PIR: A32693; A32693.
 DR HSSP: P19793; 2NLL.
 DR TRANSFAC: T02741; -
 DR FLYBASE: FBgn0003651; svp.
 DR InterPro: IPR000536; Hormone_rec_11g.
 DR InterPro: IPR001723; Stihm_receptor.
 DR InterPro: IPR001628; ZnF_C4steroid.
 DR Pfam: PF00104; hormone_rec.1.
 DR Pfam: PF00105; zf-C4; 1.
 DR PRINTS: PR00398; STRDHOMONER.
 DR PRINTS: PR00047; STROIDFINGER.
 DR PRODOM: PD000035; ZnF_C4steroid; 1.
 DR SMART: SM00430; HOL1; 1.
 DR SMART: SM00399; ZnF_C4; 1.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 DR Receptor: Transcription regulation: DNA-binding; Nuclear protein;
 KW Zinc-finger; Vision; Alternative splicing.
 KM Zinc-finger; Vision; Alternative splicing.
 FT DNA_BIND 200 265 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 200 220 C4-TYPE.
 FT ZN_FING 236 260 C4-TYPE.
 SQ SEQUENCE 543 AA; 57987 MW; 0BC189DCFA27213 CRC64;

Query Match 10.9%; Score 70; DB 1; Length 543;
 Best Local Similarity 26.7%; Pred. No. 5.3;
 Matches 27; Conservative 12; Mismatches 28; Indels 34; Gaps 5;

OY 21 SSSLNPG-----VARGHRRGQASRRMLQEGQCECKDWFLLAPRRKFTVSG 69
 DB 183 SSSNSGSGIDSKONTECVCGDKSGKHGYGPTCEG-----CKSPFKSVRRN-LTYS 236
 OY 70 LPRKQCPDHFKNVKKTRHRRHRRPNKHSRACQFLKOC 110
 DB 237 RGSRNCPID-----QHRRNQCQYCR-----LKKC 260

RESULT 9

7UP2_DROME STANDARD; PRT; 746 AA.
 ID 7UP2_DROME
 AC P16376;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Steroid receptor seven-up type 2.
 GN SVP OR NR2F3.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_Taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-90124631; PubMed-2105166;
 RA Mlodzik M., Hirrlinger Y., Weber U., Goodman C.S., Rubin G.M.;
 RT "The Drosophila seven-up gene, a member of the steroid receptor gene
 superfamily, controls photoreceptor cell fates.";
 RL Cell 60:211-224 (1990).
 CC -1- FUNCTION: PUTATIVE RECEPTOR THAT IS REQUIRED IN PHOTORECEPTOR
 CC CELLS PRECURSORS DURING EYE DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
 CC (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING AND ONLY DIFFER
 CC IN THEIR C-TERMINAL PART.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN A SUBSET OF NEURAL PRECURSORS.
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
 CC NR2 SUBFAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).

CC EMBL: M2864; AAA03014.1; -
 DR PIR: B32693; B32693.
 DR HSSP: P19793; 2NLL.
 DR FLYBASE: FBgn0003651; svp.
 DR InterPro: IPR000536; Hormone_rec_11g.
 DR InterPro: IPR001628; ZnF_C4steroid.
 DR Pfam: PF00104; hormone_rec.1.
 DR Pfam: PF00105; zf-C4; 1.
 DR PRINTS: PR00047; STROIDFINGER.
 DR PRODOM: PD000035; ZnF_C4steroid; 1.
 DR SMART: SM00430; HOL1; 1.
 DR SMART: SM00399; ZnF_C4; 1.
 DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
 DR Receptor: Transcription regulation: DNA-binding; Nuclear protein;
 KW Zinc-finger; Vision; Alternative splicing.
 KM Zinc-finger; Vision; Alternative splicing.
 FT DNA_BIND 200 265 NUCLEAR RECEPTOR-TYPE.
 FT ZN_FING 200 220 C4-TYPE.
 FT ZN_FING 236 260 C4-TYPE.
 SQ SEQUENCE 746 AA; 76830 MW; 7F256AFDA165326D CRC64;

Query Match 10.9%; Score 70; DB 1; Length 746;
 Best Local Similarity 26.7%; Pred. No. 7.2;
 Matches 27; Conservative 12; Mismatches 28; Indels 34; Gaps 5;

OY 21 SSSLNPG-----VARGHRRGQASRRMLQEGQCECKDWFLLAPRRKFTVSG 69
 DB 183 SSSNSGSGIDSKONTECVCGDKSGKHGYGPTCEG-----CKSPFKSVRRN-LTYS 236
 OY 70 LPRKQCPDHFKNVKKTRHRRHRRPNKHSRACQFLKOC 110
 DB 237 RGSRNCPID-----QHRRNQCQYCR-----LKKC 260

RESULT 10

TRBM_HUMAN STANDARD; PRT; 575 AA.
 ID TRBM_HUMAN
 AC P07204;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Thrombomodulin precursor (Fetomodulin) (TM) (CD141 antigen).
 GN THBD OR THBM.
 GN Homo sapiens (Human).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-88004395; PubMed-2820710;
 RA Suzuki K., Kusumoto H., Devashiki Y., Nishiohara J., Maruyama I.,
 RA Suzuki M., Kawahara S., Honda G., Yamamoto S., Horiguchi S.;
 RT "Structure and expression of human thrombomodulin, a thrombin
 RT receptor on endothelium acting as a cofactor for protein C
 RT activation.";
 RL EMBO J. 6:1891-1897 (1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-88024950; PubMed-2822087;
 RA Wen D., Dittman W.A., Ye R.D., Deaven L.L., Majerus P.W., Sadler J.E.;
 RT "Human thrombomodulin: complete cDNA sequence and chromosome
 RT localization of the gene.";
 RL Biochemistry 26:4350-4357 (1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-87317665; PubMed-2819876;
 RA Jackman R.W., Beeler D.L., Fritze L., Soff G., Rosenberg R.D.;
 RT "Human thrombomodulin gene is intron depleted: nucleic acid sequences
 RT of the cDNA and gene predict protein structure and suggest sites of
 RT regulatory control.";
 RL Proc. Natl. Acad. Sci. U.S.A. 84:6425-6429 (1987).
 RN [4]
 RP SEQUENCE FROM N.A.

RX MEDLINE-88227901; PubMed-2836377;
 RA Shiral T., Shiojiri S., Ito H., Yamamoto S., Kusumoto H.,
 RA Deyashiki Y., Maruyama I., Suzuki K.;
 RT "Gene structure of human thrombomodulin, a cofactor for thrombin-
 RT catalyzed activation of protein C."; *Biochem. J.* 103:281-285(1988).
 RL J. Biochem. 103:281-285(1988).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21638749; PubMed-11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.D., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.E., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Grafham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Levensalho M.H., Leverisa M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., Macdonald L.J., McEay K., McMurtry A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkhen S., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilming L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20."; *Nature* 414:865-871(2001).
 RL Nature 414:865-871(2001).
 RN [6]
 RP CARBOHYDRATE-LINKAGE SITE SER-492, AND MUTAGENESIS.
 RX MEDLINE-94029900; PubMed-8216207;
 RA Gerleitz B., Hassell T., Vlahos C.J., Parkinson J.F., Bang N.U.,
 RA Grinnell B.W.;
 RT "Identification of the predominant glycosaminoglycan-attachment site
 RT in soluble recombinant human thrombomodulin: potential regulation of
 RT functionality by glycosyltransferase competition for serine474."; *Biochem. J.* 235:131-140(1993).
 RL Biochem. J. 235:131-140(1993).
 RN [7]
 RP STRUCTURE BY NMR OF 389-407.
 RX MEDLINE-96007474; PubMed-7559494;
 RA Adler M., Seto M.H., Nitecki D.E., Lin J.H., Light D.R., Morser J.;
 RT "The structure of a 19-residue fragment from the C-loop of the fourth
 RT epidermal growth factor-like domain of thrombomodulin."; *J. Biol. Chem.* 270:23366-23372(1995).
 RL J. Biol. Chem. 270:23366-23372(1995).
 RN [8]
 RP STRUCTURE BY NMR OF 364-407.
 RX MEDLINE-96100636; PubMed-8528067;
 RA Melinger D.P., Hunter M.J., Komives E.A.;
 RT "Synthesis, activity, and preliminary structure of the fourth
 RT EGF-like domain of thrombomodulin."; *Protein Sci.* 4:1683-1695(1995).
 RL Protein Sci. 4:1683-1695(1995).
 RN [9]
 RP STRUCTURE BY NMR OF 427-444.
 RX MEDLINE-95034791; PubMed-7947766;
 RA Srinivasan J., Hu S., Hrabal R., Zhu Y., Komives E.A., Ni F.;
 RT "Thrombin-bound structure of an EGF subdomain from human
 RT thrombomodulin determined by transferred nuclear Overhauser
 RT effects."; *Biochemistry* 33:13553-13560(1994).
 RL Biochemistry 33:13553-13560(1994).
 RN [10]
 RP STRUCTURE BY NMR OF 427-444.
 RX MEDLINE-96276211; PubMed-8745396;
 RA Hrabal R., Komives E.A., Ni F.;
 RT "Structural resiliency of an EGF-like subdomain bound to its target
 RT protein, thrombin.";

RL Protein Sci. 5:195-203(1996).
 RN [11]
 RP STRUCTURE BY NMR OF 405-444.
 RX MEDLINE-98035729; PubMed-9367781;
 RA Sampoli Bentz B.A., Hunter M.J., Melinger D.P., Komives E.A.;
 RT "Structure of the fifth EGF-like domain of thrombomodulin: an
 RT EGF-like domain with a novel disulfide-bonding pattern."; *J. Mol. Biol.* 273:913-926(1997).
 RL J. Mol. Biol. 273:913-926(1997).
 RN [12]
 RP VARIANT TED TYR-486.
 RX MEDLINE-9511115; PubMed-7811989;
 RA Oehlin A.-K., Marlar R.A.;
 RT "The first mutation identified in the thrombomodulin gene in a
 RT 45-year-old man presenting with thromboembolic disease."; *Blood* 85:330-336(1995).
 RL Blood 85:330-336(1995).
 RN [13]
 RP VARIANT TED Y-486, AND VARIANTS T-43; A-79; S-495 AND L-501.
 RX MEDLINE-97341986; PubMed-9198186;
 RA Oehlin A.-K., Norlund L., Marlar R.A.;
 RT "Thrombomodulin gene variations and thromboembolic disease."; *Thromb. Haemost.* 78:396-400(1997).
 RL Thromb. Haemost. 78:396-400(1997).
 RN [14]
 RP VARIANT VAL-473.
 RX MEDLINE-97206518; PubMed-9157575;
 RA Norlund L., Holm J., Zoller B., Oehlin A.-K.;
 RT "A common thrombomodulin amino acid dimorphism is associated with
 RT myocardial infarction."; *Thromb. Haemost.* 77:248-251(1997).
 RL Thromb. Haemost. 77:248-251(1997).
 RN [15]
 RP VARIANT THR-43.
 RX MEDLINE-99057299; PubMed-9843165;
 RA Doggen C.J.M., Kunz G., Rosendaal F.R., Lane D.A., Vos H.L.,
 RA Stubbs P.J., Manger Cate V., Ireland H.;
 RT "A mutation in the thrombomodulin gene, 127G to A coding for Ala25Thr,
 RT and the risk of myocardial infarction in men."; *Thromb. Haemost.* 80:743-748(1998).
 RL Thromb. Haemost. 80:743-748(1998).
 RN [16]
 RP VARIANT VAL-473.
 RX MEDLINE-21143723; PubMed-11245641.
 RA Wu K.K., Aleksic N., Ahn C., Boerwinkle E., Folsom A.R.,
 RA Juneja H.;
 RT "Thrombomodulin Ala455Val polymorphism and risk of coronary heart
 RT disease."; *Circulation* 103:1386-1389(2001).
 RL Circulation 103:1386-1389(2001).
 CC -1- FUNCTION: THROMBOMODULIN IS A SPECIFIC ENDOTHELIAL CELL RECEPTOR
 CC THAT FORMS A 1:1 STOICHIOMETRIC COMPLEX WITH THROMBIN. THIS
 CC COMPLEX IS RESPONSIBLE FOR THE CONVERSION OF PROTEIN C TO THE
 CC ACTIVATED PROTEIN C (PROTEIN CA). ONCE EVOLVED, PROTEIN CA
 CC SCISSONS THE ACTIVATED COFACTORS OF THE COAGULATION MECHANISM,
 CC FACTOR VA AND FACTOR VIIIA, AND THEREBY REDUCES THE AMOUNT OF
 CC THROMBIN GENERATED.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: ENDOTHELIAL CELLS ARE UNIQUE IN SYNTHESIZING
 CC THROMBOMODULIN.
 CC -1- POLYMORPHISM: VARIATIONS IN THBD ARE ASSOCIATED WITH AN INCREASED
 CC RISK OF DEVELOPING THROMBOEMBOLIC DISEASES (TED).
 CC -1- DISEASE: DEFECTS IN THBD COULD BE THE CAUSE OF INHERITED TED,
 CC ALSO CALLED INHERITED THROMBOPHILIA. PATIENTS WITH TED HAVE
 CC DEFECTS OF THE HAEMOPHILIC SYSTEM WHICH CREATES A TENDENCY TO THE
 CC OCCURRENCE OF THROMBOSIS. TED PLAY AN IMPORTANT ROLE IN THE
 CC PATHOGENESIS OF VARIOUS CARDIOVASCULAR DISORDERS.
 CC -1- SIMILARITY: CONTAINS 6 EGF-LIKE DOMAINS.
 CC -1- DATABASE: NAME-PROV; NOTE-CD guide CD141 entry;
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd141.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

DR EMBL: X05495; CAA29045.1; -
DR EMBL: M16552; AAB59508.1; -
DR EMBL: J02973; AAB61175.1; -
DR EMBL: D00210; BAA00149.1; -
DR EMBL: A1049651; CAB51954.1; -
DR PIR: A27073; A27073.
DR PIR: A28307; A28307.
DR PIR: A29680; A29680.
DR PDB: 1EGT; 15-NOV-95.
DR PDB: 1FGD; 20-JUN-96.
DR PDB: 1PGE; 20-JUN-96.
DR PDB: 1TMR; 08-JUN-95.
DR PDB: 1ZAO; 29-JAN-96.
DR PDB: 1ADX; 24-DEC-97.
DR PDB: 2ADY; 24-DEC-97.
DR GLYCOSULEDB: P07204; -.
DR Genev; HGNC:11784; THBD.

Query Match 10.8%; Score 69.5; DB 1; Length 575;
Best Local Similarity 23.5%; Pred. No. 6.3;
Matches 24; Conservative 11; Mismatches 26; Indels 41; Gaps 5;

QY 5 ISSLLLLPLMTMSVSSSLNPGVARGHRDQASRW---LQEGCECECKDWFLLAPR 61
DB 209 VSSSAAVAPLGLDLMCTAP--PAVVOGHMAR-EAPGAMDCSVENGCEHACN----- 257
QY 62 RKFMTVSGLPKRCPCP-----CDHF 80
DB 258 ----AIPGAPRCQCPAGAAIQADGRCTASATQSCNDLCHEF 295

RESULT 11
SIX3.CHICK STANDARD; PRT; 314 AA.
AC 042406;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Homeobox protein SIX3 (Sine oculis homeobox homolog 3) (CSIX3).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_Taxid=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98168856; PubMed=9510037;
RA Bovolenta P., Mallamaci A., Pueilles L., Boncinelli E.;
RT "Expression pattern of csix3, a member of the six/sine oculis family
of transcription factors.";
RL Mech. Dev. 70:201-203(1998).
CC -1- FUNCTION: MAY BE INVOLVED IN VISUAL SYSTEM DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -1- SIMILARITY: BELONGS TO THE SIX/SINE OCULIS HOMEBOX FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y15106; CAA75380.1; -
DR HSSP: P41778; 1DU6.
DR TRANSFAC: TP03482; -.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR Prodom: PD000010; Homeobox; 1.
DR SMART: SM00389; HOX; 1.
DR PROSITE: PS00027; HOMEBOX_1; FALSE_NEG.
DR PROSITE: PS50071; HOMEBOX_2; 1.

```

```

KW Developmental protein; Homeobox; DNA-binding; Nuclear protein.
FT DOMAIN 44 50 POLY-GLY.
FT DNA_BIND 188 247 HOMEBOX.
FT DOMAIN 245 248 POLY-ALA.
SQ SEQUENCE 314 AA; 34677 MW; D9A04530185BA75F CRC64;

Query Match 10.7%; Score 69; DB 1; Length 314;
Best Local Similarity 26.9%; Pred. No. 4;
Matches 25; Conservative 13; Mismatches 33; Indels 22; Gaps 4;

QY 41 RMLQEGCECEC-----KDFLLAP-----RKFMTVSGLPKRCPCDHFKNV 85
DB 183 RTIWGEQKTHCFKERTSLREWYLDPPNPSKRELAQATGLTPQV-----GNWF 236
QY 86 KTRHQRHRRKPNKSRACQOFLKOCOLRSFALP 118
DB 237 KNRQRDRAAAAK-NRLQHQAIQSGHRSIAEP 268

RESULT 12
VDR.MOUSE STANDARD; PRT; 422 AA.
ID VDR.MOUSE
AC P48281;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Vitamin D3 receptor (VDR) (1,25-dihydroxyvitamin D3 receptor).
GN VDR OR NR1H1.
OS Mus musculus (mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95137405; PubMed=7835717;
RA Kamei Y., Kawada T., Fukuwatari T., Ono T., Kato S., Sugimoto E.;
RT "Cloning and sequencing of the gene encoding the mouse vitamin D
receptor.";
RL Gene 152:281-282(1995).
CC -1- FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING
THE EXPRESSION OF HORMONE SENSITIVE GENES.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC NRI SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D31969; BAA06737.1; -
DR HSSP: P03372; 1HCO.
DR TRANSFAC: T00883; -.
DR MGD: MG1103076; VDR.
DR InterPro: IPR000536; Hormone_rec_11q.
DR InterPro: IPR001628; ZnF_C4steroid.
DR Pfam: PF00104; hormone_rec; 1.
DR Pfam: PF00105; zf-C4; 1.
DR PRINTS: PRO0047; STROIDFINGER.
DR Prodom: PD000035; ZnF_C4steroid; 1.
DR SMART: SM00430; HOL1; 1.
DR SMART: SM00399; ZnF_C4; 1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
DR Receptor: Transcription regulation; DNA-binding; Nuclear protein;
KW Zinc-finger; Phosphorylation.
FT DNA_BIND 24 89 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 24 44 C4-TYPE.
FT ZN_FING 60 84 C4-TYPE.

```

```

FT DOMAIN 90 186 HINGE.
FT ZN_FING 187 422 LIGAND-BINDING.
SQ SEQUENCE 422 AA; 47851 MW; 4704CG8172445732 CRC64;
Query Match 10.6%; Score 68.5; DB 1; Length 422;
Best Local Similarity 30.1%; Pred. No. 6;
Matches 22; Conservative 13; Mismatches 15; Indels 23; Gaps 6;

OY 50 CE-CKDWFLAPRRKMTVSGLPKKQPCDHFKNVKTTHORHHRKPNKHSRACQQLK 108
DB 41 CEGCKGFRRSRMKRAKFT-----CP---FNGDCRITKDNR-----RHQSCR--LK 82
OY 109 QC-----QLRSFAL 117
DB 83 RCVDIGMKKEFIL 95

RESULT 13
VDR_XENLA STANDARD; PRT; 422 AA.
AC 013124;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Vitamin D3 receptor (VDR) (1,25-dihydroxyvitamin D3 receptor).
GN VDR OR NR11.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_Taxid=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RA MEDLINE=97307679; PubMed=9165021;
RA Li Y.C., Bergwitz C., Jueppner H., Demay M.B.;
RT "Cloning and characterization of the vitamin D receptor from Xenopus
laevis."
RL Endocrinology 138:2347-2353(1997).
CC -1- FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING
THE EXPRESSION OF HORMONE SENSITIVE GENES.
CC -1- SURCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: DETECTED IN ALL THE TISSUE EXAMINED. HIGHEST
LEVEL IN SMALL INTESTINE AND SKIN.
CC -1- DEVELOPMENTAL STAGE: FIRST DETECTED AT STAGE 13. INCREASES
GRADUALLY AND PEAKS AT STAGE 57-61 THEN DECREASES TO THE LEVEL
SEEN IN ADULT.
CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC -1- SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; U91846; AAB58585.1; -
DR HSSP; O75469; 11LG.
DR InterPro; IPR000536; Hormone_rec_119.
DR Pfam; PF00104; hormone_rec.1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00047; STROIDFINGER.
DR ProDom; PD000035; Znf_C4steroid.1.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00339; Znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
Zinc-finger.

```

```

FT DNA_BIND 25 90 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 25 45 C4-TYPE.
FT ZN_FING 61 85 C4-TYPE.
FT DOMAIN 91 188 HINGE.
SQ SEQUENCE 422 AA; 48188 MW; C8A9F25414FE9D5 CRC64;
Query Match 10.6%; Score 68.5; DB 1; Length 422;
Best Local Similarity 30.1%; Pred. No. 6;
Matches 22; Conservative 13; Mismatches 15; Indels 23; Gaps 6;

OY 50 CE-CKDWFLAPRRKMTVSGLPKKQPCDHFKNVKTTHORHHRKPNKHSRACQQLK 108
DB 42 CEGCKGFRRSRMKRAKFT-----CP---FNGDCRITKDNR-----RHQSCR--LK 83
OY 109 QC-----QLRSFAL 117
DB 84 RCVDIGMKKEFIL 96

RESULT 14
VDR_RAT STANDARD; PRT; 423 AA.
AC P13053;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Vitamin D3 receptor (VDR) (1,25-dihydroxyvitamin D3 receptor).
GN VDR OR NR11.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=89071726; PubMed=2849110;
RA Burmester J.K., Wiese R.U., Maeda N., Deluca H.F.;
RT "Structure and regulation of the rat 1,25-dihydroxyvitamin D3
receptor."
RL Proc. Natl. Acad. Sci. U.S.A. 85:9499-9502(1988).
RN [2]
CC -1- FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING
THE EXPRESSION OF HORMONE SENSITIVE GENES.
CC -1- SURCELLULAR LOCATION: Nuclear.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC -1- SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
DR EMBL; J04147; AAA41089.1; -
DR PIR; A31761; A31761.
DR PIR; A31367; A31367.
DR HSSP; P03372; 1HCQ.
DR TRANSFAC; T00862; -
DR InterPro; IPR000536; Hormone_rec_119.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; hormone_rec.1.
DR Pfam; PF00105; zf-C4; 1.
DR PRINTS; PR00047; STROIDFINGER.

```

```

DR Prodom; PD000035; znf_C4steroid; 1.
DR SMART; SM00430; HOLI; 1.
DR SMART; SM00399; znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR RECEPTOR; 1.
KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KM Zinc-finger; Phosphorylation.
FT DNA_BIND 24 89 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 24 44 C4-TYPE.
FT ZN_FING 60 84 C4-TYPE.
FT DOMAIN 90 187 HINGE.
FT DOMAIN 188 423 LIGAND-BINDING.
SQ SEQUENCE 423 AA; 47813 MW; 1A0E519A9DCCBE90 CRC64;

Query Match 10.6%; Score 68.5; DB 1; Length 423;
Best Local Similarity 30.1%; Pred. No. 6;
Matches 22; Conservative 13; Mismatches 15; Indels 23; Gaps 6;

QY 50 CE-CKDMFLAPRRKFTVSGLPKQCPDHFNGVAKTRHQRHRRPNKHSRACQFLK 108
DB 41 CECKGFFRRSMRKALFT-----CP---FNGDCRITRDNR-----RHQACR--LK 82

QY 109 QC---QLRSFAL 117
DB 83 RCVDIGMKKEFIL 95

RESULT 15
VDR_BOVIN STANDARD; PRT; 424 AA.
AC 028037;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Vitamin D3 receptor (VDR) (1,25-dihydroxyvitamin D3 receptor).
GN VDR OR NR11.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97034797; PubMed=8880453;
RA Nelberrgs H.L., Bosworth B.T., Reinhardt T.A.;
RT "Nucleotide sequence of the bovine vitamin D3 receptor.";
RL J. Dairy Sci. 79:1313-1315(1996).
CC -!- FUNCTION: VDR MEDIATES THE ACTION OF VITAMIN D3 BY CONTROLLING
CC THE EXPRESSION OF HORMONE SENSITIVE GENES.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- DOMAIN: COMPOSED OF THREE DOMAINS: A MODULATING N-TERMINAL DOMAIN,
CC A DNA-BINDING DOMAIN AND A C-TERMINAL STEROID-BINDING DOMAIN.
CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC NRI SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U50200; AAB01543.1; -.
DR HSSP; 075469; IILG.
DR InterPro; IPR000536; Hormone_rec_11g.
DR InterPro; IPR001628; znf_C4steroid.
DR Pfam; PF00104; hormone_rec; 1.
DR PRINTS; PR00047; STROIDFINGER.
DR PRODOM; PD000035; znf_C4steroid; 1.
DR SMART; SM00430; HOLI; 1.
DR SMART; SM00399; znf_C4; 1.
DR PROSITE; PS00031; NUCLEAR RECEPTOR; 1.

```

```

KW Receptor; Transcription regulation; DNA-binding; Nuclear protein;
KM Zinc-finger; Phosphorylation.
FT DNA_BIND 21 86 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 21 41 C4-TYPE.
FT ZN_FING 57 81 C4-TYPE.
FT DOMAIN 87 188 HINGE.
FT DOMAIN 189 424 LIGAND-BINDING.
SQ SEQUENCE 424 AA; 47957 MW; E9E24926CE38C87D CRC64;

Query Match 10.6%; Score 68.5; DB 1; Length 424;
Best Local Similarity 30.1%; Pred. No. 6;
Matches 22; Conservative 13; Mismatches 15; Indels 23; Gaps 6;

QY 50 CE-CKDMFLAPRRKFTVSGLPKQCPDHFNGVAKTRHQRHRRPNKHSRACQFLK 108
DB 38 CECKGFFRRSMRKALFT-----CP---FNGDCRITRDNR-----RHQACR--LK 79

QY 109 QC---QLRSFAL 117
DB 80 RCVDIGMKKEFIL 92

```

Search completed: April 16, 2003, 12:15:33
 Job time : 16 secs

